

Overview of Sentiment Analysis on Social Media

S.P.Ramya

Computer Science and Engineering
Periyar Maniammai Institute of Science and Technology,
Thanjavur, Tamilnadu

D.Nivetha

Computer Science and Engineering
Periyar Maniammai Institute of Science and Technology,
Thanjavur, Tamilnadu

Abstract— Analysis of public information from social media could yield interesting results and insights into the world of public opinions about almost any product, service or personality. Social network data is one of the most effective and accurate indicators of public sentiment. As a result there has been an eruption of interest in people to mine these vast resources of data for opinions. Sentiment Analysis or Opinion Mining is the computational treatment of opinions, sentiments and subjectivity of text. In this paper we will be discussing a methodology which allows utilization and interpretation of social media data to determine public opinions. Developing a program for sentiment analysis is an approach to be used to computationally measure customers' perceptions. This paper reports on the design of a sentiment analysis, extracting and training a vast amount of tweets. Results classify customers' perspective via tweets into positive and negative, which is represented in diagrams.

Keywords—Opinion Mining; Sentiment Analysis

I. INTRODUCTION

With the explosive growth of social media (e.g. blogs, micro-blogs, forum discussions and reviews) in the last decade, the web has drastically changed to the extent that nowadays billions of people all around the globe are freely allowed to conduct many activities such as interacting, sharing, posting and manipulating contents. This enables us to be connected and interact with each other anytime without geographical boundaries, as opposed to the traditional structured data available in databases. The resulted unstructured user-generated data mandates new computational techniques from social media mining, while it provides us opportunities to study and understand individuals at unprecedented scales. Sentiment analysis is one class of computational techniques which automatically extracts and summarizes the opinions of such immense volume of data which the average human reader is unable to process.

This ocean of opinionated postings in social media is central to the individuals' activities as they impact our behaviors and help reshape businesses.

Nowadays, not only individuals are no longer limited to asking friends and family about products but also businesses, organizations and companies do not require conducting surveys or polls for opinions about products, as there are tons of user reviews and discussions in public forums on the Web.

There are thus numerous immediate and practical applications and industrial interests of collecting and studying such opinions by using computational sentiment analysis techniques, spreading from consumer products, services, healthcare, and financial services to social events, political elections and more recently crisis management and natural disasters.

Social media has pervasively played an increasing role and they have become an important alternative information channel to traditional media in the last five years during emergencies and disasters, where they rank as the fourth most popular sources to access necessary information during emergencies. In 2005 for instance, when Hurricane Katrina slammed the U.S. gulf coast, there was no Twitter for news update while Facebook was not that much famous yet. Compare, for example, Hurricane Katrina to the Haiti earthquake on January 2010.

During latter, people used Twitter, Facebook, Flickr, blogs and YouTube to post their experience in form of texts, photos and videos during the earthquake which resulted in donating 8 million U.S. dollars to the Red Cross which vividly demonstrates the power of social media in propagating information during emergencies. Social media could be actually leveraged to keep the problem informed, help locate loved ones, and express support or notify authorities during emergencies and disasters.

Sentiment analysis of disaster related posts in social media could help to detect posts that contribute to the situational awareness and better understand the dynamics of the network including user's feelings, panics and concerns by identifying the polarity of sentiments expressed by users during disaster events to improve decision making.

Sentiment information could also be used to project the information regarding the devastation and recovery situation and donation requests to the crowd in better ways.

Interactive tools such as visual analytic methods could help us to make a large amount of complex information more readable and interpretable, if integrated by computational approaches, as the effectiveness of most computational techniques is limited due to several factors. Interactive visual analytics provide intuitive ways of making sense of large amount of posts available in social media. These techniques are now widely used in social media data and contribute in many areas of exploratory data analysis.

Despite most social media visualization approaches which rely solely on geographical and temporal features, there are some systems which are able to exploit the sentiments of the data such which help improving visualization.

By the end of this chapter, the reader is expected to learn about the sentiment analysis and disaster management concepts as well as the state-of-the-art approaches and the applications of visual analytics in these contexts.

II. SENTIMENT ANALYSIS

Sentiment analysis (a.k.a sentiment classification, opinion mining, subjectivity analysis, polarity classification, affect

analysis, etc.) is the multidisciplinary field of study that deals with analyzing people's sentiments, attitudes, emotions and opinions about different entities such as products, services, individuals, companies, organizations, events and topics and includes multiple fields such as natural language processing (NLP), computational linguistics, information retrieval, machine learning and artificial intelligence. It is set of computational and NLP based techniques which could be leveraged in order to extract subjective information in a given text unlike factual information, opinions and sentiments are subjective.

Despite the recent surge of interest in sentiment analysis since the term was coined by Nasukawa et al. in 2003, the demand for information on sentiment and opinion during decision-making situations dates back to long before the widespread use of the World Wide Web. Opinions are central to almost all human activities as they could influence our behaviors specially when making a decision. For example, many of us may have asked their friends to recommend a dishwasher or to explain who they might vote for during elections, or even requested reference letters from colleagues regarding job applications.

Now, opinions and experiences of numerous people that are neither our acquaintances nor professional critics are readily available thanks the Internet and the Web. This is not limited to individuals only; businesses, organizations and companies are also eager to know consumer's opinions about their products and services. In the past, when a business needed consumer opinions, it conducted surveys and opinion polls. Nowadays, one is no longer limited to asking friends and family or conducting surveys for opinions about products; instead one can use volumes of user reviews and discussions in public forums on the Web. Indeed, the Web has dramatically changed the way that people express their opinions about products, services, companies, individuals and social events.

There are now many Internet forums, discussion groups, blogs and even micro-blogs that are well suited for the users to freely post reviews about products and express their views on almost anything online. These users-generated contents and word-of-mouth behavior are sources of information with many immediate and practical applications.

a) Document level:

The aim here is to determine the overall sentiment of an entire document. For example given a product review, the task is to determine whether it expresses positive or negative opinions about the product. This level looks at the document as a single entity, thus it is not extensible to multiple documents.

b) Sentence level:

This level of analysis is very close to subjectivity classification and the task at this level is limited to the sentences and their expressed opinions. Specifically, this level determines whether each sentence expresses a positive, negative or neutral opinion.

c) Entity and aspect level:

Instead of solely analyzing language constructs (e.g. documents, paragraphs, sentences), this level (a.k.a feature level) provides finer-grained analysis for each aspect (or feature) i.e., it directly looks at the opinions for different aspects itself. The aspect-level is more challenging than both

document and sentence levels and consists of several sub-problems. It finds different available sentiment.

Sentiment analysis methods could be categorized into two groups, language processing based and application oriented methods. We describe the state-of-the-art approaches in each category and highlight their contributions. Then we conclude this section with a brief overview on visual analytics approaches in sentiment analysis.

III. LOGICAL COMPONENTS

The Sentiment and Knowledge Mining system used in this study is built on the following components:

1. A Crawler, an adaptive and selective component that gathers documents from Internet/Intranet or Database sources.
2. A Semantic Engine, which identifies relevant knowledge in the texts, by detecting semantic relations and facts.
3. A Search Engine that enables Natural Language, Semantic and Semantic-Role queries.
4. A Machine Translation Engine, which enables automatic translation of search results.
5. A Geo-referentiation Engine, which enables an interactive geographical representation of documents.
6. A Classification Engine which classifies search results into clusters and sub-clusters recursively, highlighting meaningful relationships among them, or assigns documents to predefined thematic groups.

IV. SENTIMENT VISUALIZATION

Visual analytics focuses on providing an intuitive way of making sense of large amount of posts available in social media. It is widely used in social media data and contributes in many areas of exploratory data analysis, such as geographical analysis, information diffusion and business prediction. While most social media visualization approaches rely on geographical and temporal features, some systems exploit the semantic of the data such as sentiments to improve visualization. For example, there exist some Websites that provide mash up applications to visualize and analyze tweets, including TrendsMap3, Twitalyzer4, and Geotwitterous5, some of which provide sentiment analysis as well. In the remainder of this section, we explore some of the existing systems and tools that are able to visualize sentiment. IN-SPIRE is a visual analytic tool for blog analysis which helps users to harvest blogs and classify them with respect to their contents. IN-SPIRE supports sentiment visualization of blogs and streaming contents. The sentiment of lexicons is categorized into positive/negative, virtue/vice, power coop/conflict and pleasure/pain according to Gregory et al. and is visualizing as a rose plot. Each pair of category is shown with a different shade of a same color. Moreover, the sizes of petals are different in terms of the amount of sentiment.

V. CONCLUSION

This paper describes a Sentiment Analysis study performed on over than 1000 Facebook posts about newscasts, comparing the sentiment for Rai - the Italian public broadcasting service - towards the emerging and more dynamic private company La7. It maps Sentiment Analysis on Social Media with observations and measurable data. Its results accurately reflect the reality high lightening the importance of Facebook as a

platform for online marketing. Monitoring the social media activities is a good way to measure customers' loyalty and interests, keeping track of their sentiment towards brands or products. This study has been performed by a Knowledge Mining system used by some security sector-related government institutions and agencies in Italy to limit information overload in OSINT and Web Mining. The linguistic and semantic approaches implemented in this system enable the research, the analysis, the classification of great volumes of heterogeneous documents, helping documental analysts to cut through the information labyrinth, analysts to take account of complexity of public views, assigning automatically a sentiment polarity, rapidly accessing all the potential texts of interest.

REFERENCES

- [1] Reza Zafarani, Mohammad Ali Abbasi, and Huan Liu. Social media mining: an introduction. Cambridge University Press, 2014.
- [2] Jure Leskovec, Daniel Huttenlocher, and Jon Kleinberg. Predicting positive and negative links in online social networks. In Proceedings of the 19th international conference on World wide web, pages 641–650. ACM, 2010.
- [3] G.Beigi, M.Jalili, H.Alvari, and G.Sukthankar. Leveraging community detection for accurate trust prediction. In ASE International Conference on Social Computing, June 2014.
- [4] Jiliang Tang, Xia Hu, and Huan Liu. Social recommendation: a review. Social Network Analysis and Mining, 3(4):1113–1133, 2013.
- [5] Hamidreza Alvari, Alireza Hajibagheri, and Gita Sukthankar. Community detection in dynamic social networks: A game-theoretic approach. In Advances in Social Networks Analysis and Mining (ASONAM), 2014 IEEE/ACM International Conference on, pages 101–107. IEEE, 2014.
- [6] Michaela Goetz, Jure Leskovec, Mary McGlohon, and Christos Faloutsos. Modeling blog dynamics. Citeseer, 2009.
- [7] Hamidreza Alvari, Sattar Hashemi, and Ali Hamzeh. Discovering overlapping communities in social networks: A novel game-theoretic approach. AI Communications, 26(2):161–177, 2013.
- [8] Bruce R. Lindsay. Social Media and Disasters: Current Uses, Future Options, and Policy Considerations. Technical report, Congressional Research Service, September.
- [9] Alfredo Cobo, Denis Parra, and Jaime Nav'on. Identifying relevant messages in a twitter-based citizen channel for natural disaster situations. In Proceedings of the 24th International Conference on World Wide Web Companion, pages 1189–1194, 2015.
- [10] Huiji Gao, Geoffrey Barbier, and Rebecca Goolsby. Harnessing the crowdsourcing power of social media for disaster relief. IEEE Intelligent Systems, (3):10–14, 2011.